

Blair Chen

github.com/chenblair | blairc@andrew.cmu.edu | 408.533.3472

EDUCATION

Carnegie Mellon University — School of Computer Science Pittsburgh, PA

MS in Machine Learning | GPA 4.3 / 4.3 Dec 2021

Coursework: Machine Learning for Large Datasets, Probabilistic Graphical Models, Convex Optimization, Deep Learning, Deep Learning Systems (Planned), Machine Learning in Practice (Planned)

BS in Artificial Intelligence | GPA 4.0 / 4.0 Dec 2020

SKILLS

Python3 / C++ / Bash

Pytorch / Numpy / Pandas / PySpark

EXPERIENCE

TruEra Intelligence Team Redwood City, CA

Machine Learning Intern May 2021 - Present

- ❖ Developing novel global interpretability metrics and visualizations for NLP models using Integrated Gradients

Microsoft Speech Team Redmond, WA

Software Engineering Intern May 2020 - Aug 2020

- ❖ Optimized hyperparameter scheduling on a *distributed* neural network training algorithm for speech recognition
- ❖ Invented a custom scheduler written in Pytorch DDP to provide reasonable baselines on unseen datasets
- ❖ Discovered a hyperparameter scheduling strategy that improved training speed by 11%
- ❖ Designed and implemented a novel hyperparameter scheduling framework for flexible adaptive scheduling (patent pending)

Nuro Autonomy Team Mountain View, CA

Software Engineering Intern Jan 2020 - May 2020

- ❖ Built the base functionality of a prediction system for acting safely in the case of occluded vehicles
- ❖ Designed a simple heuristic controller written in C++ that led to over 50x simulation speedup
- ❖ Fine-tuned system to simulate realistic and safe driving behavior featuring deferred reaction to oncoming traffic
- ❖ Developed a Qt C++ widget that allowed playback and comparison of multiple vehicle simulations

Airbnb Content Intelligence Team San Francisco, CA

Software Engineering Intern May 2019 - Aug 2019

- ❖ Built an end-to-end system for fake account detection on an intelligence pipeline, including a backend Elixir server, a React labeler frontend, and a machine learning backend that labeled 10,000s of fake accounts a day
- ❖ Performed data analysis using Jupyter and Pandas and developed a labeler document optimizing for high-quality labels that informed over a hundred labeler-hours
- ❖ Trained a multilayer Bert-LSTM model to flag fake accounts automatically, automating hours a week of operations work

RESEARCH

CMU Machine Learning Department Sep 2019 - Present

Advisor: Dr. Louis-Philippe Morency

- ❖ Experimented with a state-of-the-art novel loss function for uncertainty training in the context of noisy labels and adversarial training compared to current methods (<https://arxiv.org/abs/2002.06541>)
- ❖ Explored the properties and applications of label smoothing during training for neural network training
- ❖ Uncovered un-interpretability of attention values for overparameterized RNN training regimes

HONORS

Phi Beta Kappa

USA Computing Olympiad (Platinum)

USA Math Olympiad